Volume 16 Number 10 1988 Nucleic Acids Research

Compositional compartmentalization and compositional patterns in the nuclear genomes of plants

Julio Salinas, Giorgio Matassi[1], Luis M.Montero and Giorgio Bernardi[1]*

Departamento de Bioquimica, Instituto Nacional de Investigaciones Agrarias, Carretera de la Coruna, 28040 Madrid, Spain and [1]Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France

Received March 3, 1988; Revised and Accepted April 21, 1988

ABSTRACT

We report here results which indicate (i) that the nuclear genomes of angiosperms is characterized by a compositional compartmentalization and an isochore structure; and (ii) that the nuclear genomes of some Gramineae exhibit strikingly different compositional patterns compared to those of many dicots. Indeed, the compositional distribution of nuclear DNA molecules (in the 50-100 Kb size range) from three dicots (pea, sunflower and tobacco) and three monocots (maize, rice and wheat) were found to be centered around lower (41%) and higher (45% for rice, 48% for maize and wheat) GC levels, respectively (and to trail towards even higher GC values in maize and wheat). Experiments on gene localization in density gradient fractions showed a remarkable compositional homogeneity in vast (>100-200 Kb) regions surrounding the genes. On the other hand, the compositional distribution of coding sequences (GenBank and literature data) from dicots (several orders) was found to be narrow, symmetrical and centered around 46% GC, that from monocots (essentially barley, maize and wheat) to be broad, asymmetrical and characterized by an upward trend towards high GC values, with the majority of sequences between 60 and 70% GC. Introns exhibited a similar compositional distribution, but lower GC levels, compared to exons from the same genes.

INTRODUCTION

At present, the molecular genetics of plants is a very rapidly expanding area of research. If we compare, however, our knowledge of nuclear genes from plants and animals, it is obvious that the former is much more limited than the latter. For example, as far as primary structures are concerned, data are presently available for only about 200 coding sequences from plants; this figure is lower than that from the human genome alone. If plant genomes are considered, the situation is not very different from that prevailing for animal genomes about twenty years ago. Indeed, the available information on plant DNAs essentially concerns buoyant density, methylation and

reassociation kinetic data for a relatively small number of species (1-3).

Here, we report results on the genomes of dicotyledons (dicots) and monocotyledons (monocots), which were obtained by using both the rationale and the experimental approaches previously developed for the vertebrate genomes (see refs. 4-8), namely by investigating the compositional distributions of both genes and DNA molecules (in the 50-100 Kb size range). At the gene level, we have used all data available in the GenBank (as of May 1987), as well as data from the literature. At the DNA level, we have used preparative centrifugation in density gradients in the presence of a sequence-specific ligand (9,10) to study DNAs from three dicots, (pea, sunflower, tobacco) and three monocots, (maize, rice and wheat).

Basically, the results obtained indicate that the genomes of the angiosperms investigated are characterized by (i) a compositional compartmentalization and an isochore organization; and by (ii) striking differences in the compositional distributions of DNA molecules, coding sequences and introns between the dicots and monocots studied.

MATERIALS AND METHODS

Preparation and fractionation of nuclear DNA

Ethiolated seedlings from three monocots , maize, *Zea mays* (single cross XL72), rice, *Oryza sativa* (cv. Bahia), wheat, *Triticum aestivum* (cv. Rivereno), and three dicots, pea, *Pisum sativum* (cv. Desso), sunflower, *Helianthus annuus* (inbred line CHS-89), tobacco, *Nicotiana tabacum* (cv. Burley), were used to prepare nuclear DNA. This was obtained by using the method of Kislev and Rubenstein (11) with minor changes. The average size of DNA molecules in all cases was comprised between 50 and 100 Kb, as determined by electrophoretic mobility. Fractionation of DNA by preparative centrifugation in Cs_2SO_4 density gradient in the presence of BAMD (3,6 bis (acetato-mercuri-methyl dioxane)) and analytical centrifugation of the preparative fractions in CsCl density gradient were carried out as previously described (9,10).

Restriction endonuclease digestion and hybridization

DNA samples from the fractions of Cs_2SO_4/BAMD preparative gradients were digested with EcoRI restriction endonuclease (Boehringer, Mannheim, FRG) using the conditions given by the supplier. Electrophoresis was carried out on 0.8% horizontal agarose gels in 8.9 mM Tris, 8.9 mM H_3BO_3, and 2.5 mM EDTA, pH 8.3. DNA was then denatured and transferred to nitrocellulose filters as described (10). Filters were pre-hybridized at 65°C for 3 hours in 3 x SSC, 0.1% SDS, 5 x Denhardt's solution and 0.1 mg/ml denatured salmon sperm DNA. Overnight hybridization was done in the pre-hybridization solution supplemented with the 32P-labeled probe (100 µg; specific radioactivity was equal to $5.0 \times 10^8 - 1.0 \times 10^9$ cpm/ug DNA). After hybridization, filters were washed in 3 x SSC, 0.1% SDS at 65°C.

Probes

The wheat high molecular weight (HMW) glutenin gene probe, obtained from Dr. J.M. Malpica, was a 1.5 Kb Hind III segment from a 3.5 Kb genomic clone (pHSB26) inserted into the Sma I/BamH I site of plasmid pUC8 (12).

The wheat chlorophyll a/b-binding protein (Cab) gene probe, obtained from Dr. N.-H. Chua, was a 1.6 Kb genomic clone (whAB1.6) inserted into the Pst I site of plasmid pUC13 (13).

Gene data

GC levels of coding sequences (from the initial AUG to the termination codon), of first + second and third positions, and of introns were obtained from GenBank Release 50 (May, 1987). The ACNUC retrieval system (14) was used. Data concerning genes not yet available in GenBank were obtained from the literature. A total of 204 genes were investigated, 116 from dicots and 88 from monocots. A listing of the coding sequences from GenBank, a reference list for the coding sequences from the literature and a list of the genes studied in both their exons and their introns will be provided upon request.

RESULTS

Compositional distribution of DNA molecules

The analytical CsCl profile of unfractionated DNAs were characterized by modal buoyant densities of 1.6969, 1.6955 and

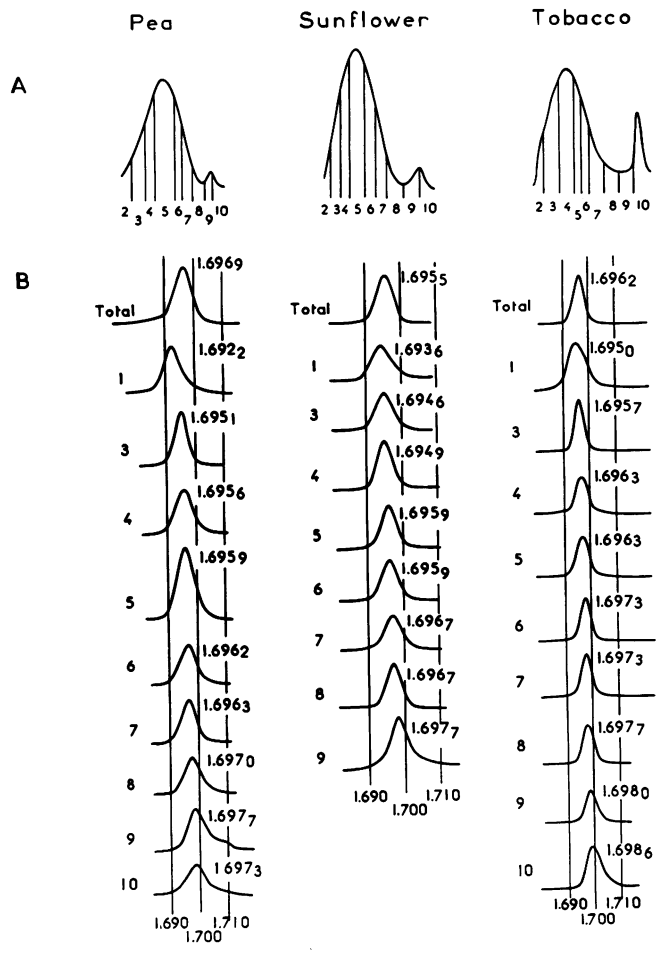

Fig. 1. Fractionation of DNAs from three dicots (pea, sunflower and tobacco) by preparative Cs_2SO_4/BAMD density gradient centrifugation. 10 A260 units of DNA were centrifuged at 30°C in 0.4 M Na_2SO_4, 20 mM $Na_2B_4O_7$, pH 9.4, 1.58 M Cs_2SO_4 for 68 hours with a BAMD to DNA molar ratio, rf, equal to 0.14, using a 50 Ti rotor at 36,000 rpm. Panel A shows the transmission profile of fractionated DNA, as recorded at 253.7 nm. Panel B shows the analytical CsCl profiles of total (unfractionated) DNA and of pooled fractions. Corresponding fractions from two preparative centrifugation tubes were pooled. The analytical CsCl profiles of fractions 2 are not shown because of their close similarity with fraction 1 (pellets) and of the very small amounts of DNA contained in them. Fraction 10 of sunflower DNA was accidentally lost in the experiment shown.

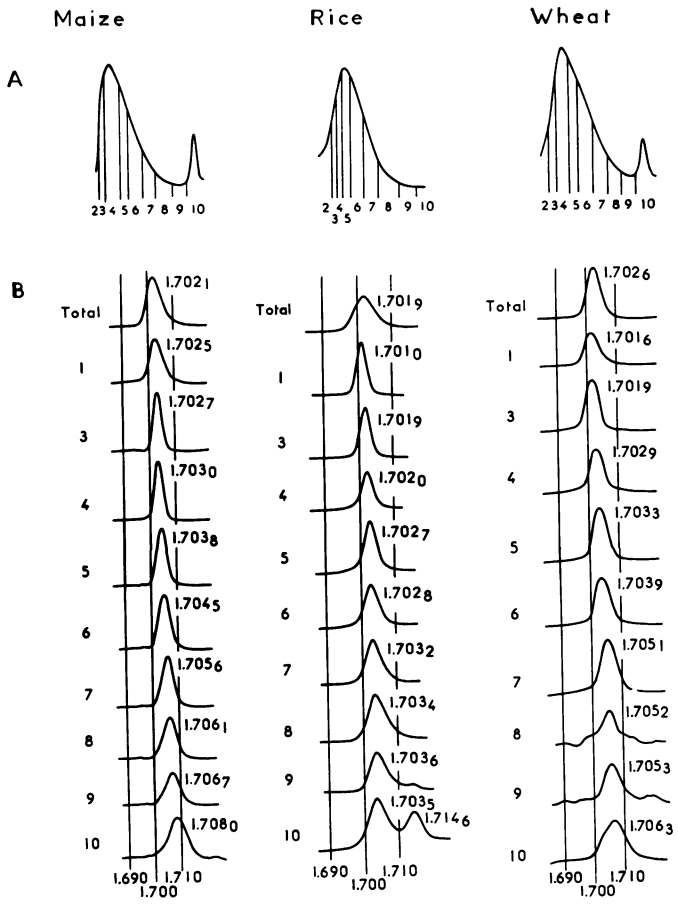

Fig. 2. Fractionation of DNAs from three monocots (maize, rice, wheat) by Cs_2SO_4/BAMD density gradient centrifugation. See legend of Fig. 1 for the experimental conditions and other indications.

1.6962 g/cm^3 for pea, sunflower and tobacco, respectively (Fig. 1), and of 1.7021, 1.7019 and 1.7026 g/cm^3 for maize, rice and wheat, respectively (Fig. 2). In order to convert the buoyant densities of plant DNAs into GC values, account was taken of their degrees of methylation. (Methylation causes a decrease in buoyant density of about 0.7 mg/cm^3 per 1% 5-methyl cytosine; 15-17). This was done using methylation data from the literature (2,17,18). After correction for methylation, the modal buoyant

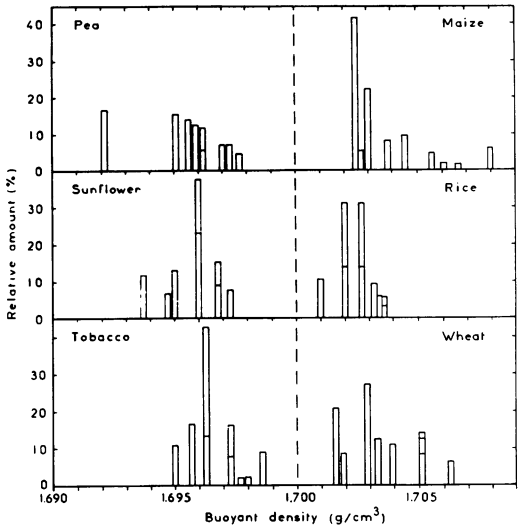

Fig. 3. Histograms showing the relative amounts and buoyant densities in CsCl of DNA fractions obtained by preparative Cs_2SO_4/BAMD density gradient centrifugation from three dicots (pea, sunflower, tobacco) and three monocots (maize, rice and wheat). The conditions used led to a very large amount of pelleted DNA in the case of maize. In another experiment (not shown), carried out at rf = 0.10 (see legend of Fig. 1), the pellet (fraction 1) only represented 12% of DNA, while fractions 2, 3 and 4 corresponded to 5%, 16% and 25% of total DNA. Buoyant densities of fractions 1-4 were 1.7013, 1.7014, 1.7018, 1.7032 g/cm^3, respectively. Horizontal lines within the vertical bars separate DNA samples having the same buoyant density but derived from different fractions. A rice satellite DNA banding at 1.7146 g/cm^3 (see Fig. 2) is not shown on the histogram.

densities of the DNAs were found to correspond to about 41% for the dicots, 45% for rice, and 48% GC for maize and wheat.

Preparative Cs_2SO_4/BAMD density gradient centrifugation was performed under conditions leading to pelleting 10-20% of DNA, the remainder being separated into nine fractions. The pellet and the fractions were used to estimate the relative amounts and the buoyant densities in CsCl of the DNA contained in them. Figs. 1 and 2 present the recordings of the preparative fractionations and the analytical CsCl profiles of the fractions.

These results allowed the construction of histograms providing some information on the compositional distribution of

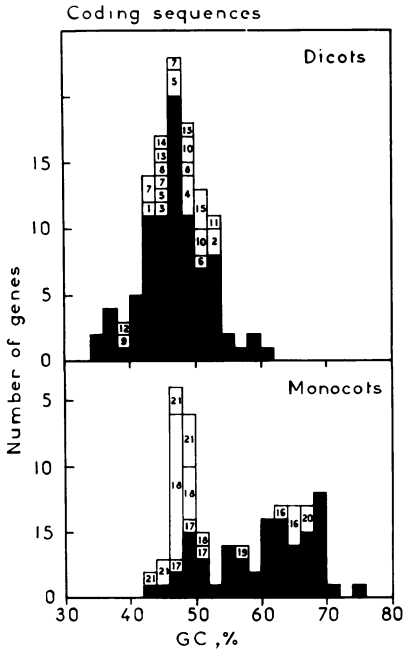

Fig. 4. The numbers of genes from dicots and monocots are plotted against the GC levels of the corresponding coding sequences; a 2% GC window was used. Genes belonging to the same multigene family were counted only once in the solid bar histogram, unless evidence was available for their non-clustered chromosomal location. The numbers and GC contents of additional genes from multigene families are represented by the open bar histogram. These were numbered as follows : 1, Peroxidase (amoracia); 2, Cab (arabidopsis); 3, Actin (soybean); 4, Glycinin (soybean); 5, Heat shock protein (soybean); 6, Rubisco (soybean); 7, Leghemoglobin (soybean); 8, Rubisco (tomato); 9, Wound-inducible proteinase inhibitor II (tomato); 10, Cab (petunia); 11, Phytohemagglutinin, Pha-L (common bean); 12, Albumin (pea); 13, Rubisco (pea); 14, Chalcone synthase (ranunculus); 15, Rubisco (petunia); 16,α-amylase (barley); 17, Hordein (barley); 18, Gliadin (wheat); 19, HMW glutenin (wheat); 20, Histone (maize); 21, Zein (maize). Storage protein genes are underlined.

DNA molecules (Fig. 3). These histograms were characterized by relatively broad GC ranges (4-6%), by the lack of overlap of dicot and monocot distributions, and by a skewness towards high GC in maize and wheat.

Compositional distribution of coding sequences

Fig. 4 displays histograms concerning the compositional

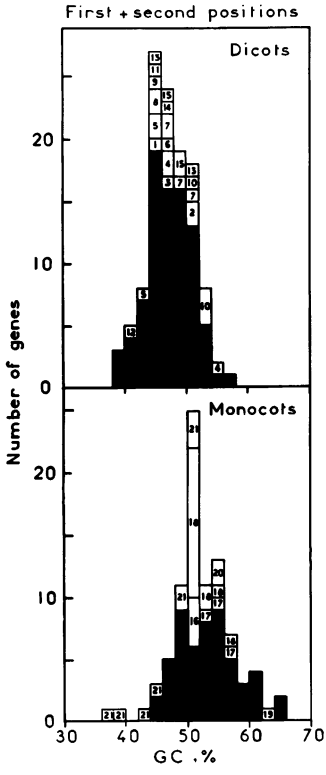

Fig. 5. The numbers of genes from dicots and monocots are plotted against the GC levels of the first + second codon positions of the corresponding coding sequences. Other indications as in Fig. 4.

distribution of coding sequences from dicots and monocots, as obtained from GenBank and from the literature. Coding sequences belonging to the same gene family were close in GC level. Unless evidence was available for their non-clustered nature, genes from the same multigene family of a given plant were counted only once in the solid bars of the histograms of Fig. 4, whereas additional copies were represented in the open bars. This was done in order to avoid biasing the histograms by taking into consideration several clustered genes having the same composition. It should be stressed, however, that histograms would only be seriously biased if the coding sequences for

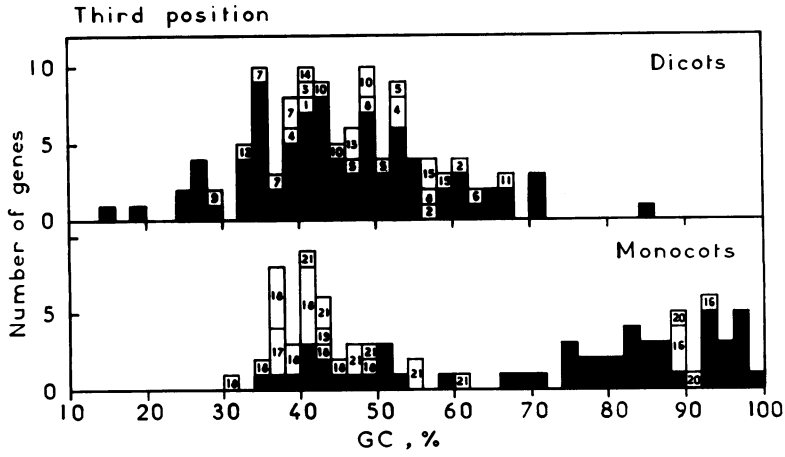

Fig. 6. The numbers of genes from dicots and monocots are plotted against the GC levels of third codon positions of the corresponding coding sequences. Other indications as in Fig. 4. In the case of dicots, the single value beyond 80% GC corresponds to the methionine-rich protein gene from Bertholletia excelsa (19).

storage proteins of monocots, which are low in GC and known to be clustered, were to be considered individually.

The compositional distribution of coding sequences from dicots was fairly symmetrical with an average GC of 46% (σ= 3.8%) and a range comprised between 34 and 62% GC. In contrast, the distribution found for coding sequences from monocots was characterized by a 42 to 76% GC range, by a strong upward increase towards high GC levels and by the majority of sequences being comprised between 60 and 70% GC (namely between values about 20% higher in GC than most coding sequences from dicots).

The compositional distributions of codon first + second positions for the genes of Fig. 4 were narrower than those of coding sequences (Fig. 5). In the case of monocots the distribution was also much less asymmetrical and centered on a value, 53.3% GC (σ= 4.8%), significantly higher than in the case of dicots, 47.2% (σ= 3.7%). The distributions of third codon positions spred over for both dicots and monocots (Fig. 6). The range covered was 15-72% in the case of dicots (with a single value beyond the upper limit; see legend of Fig. 6) and 35-100%

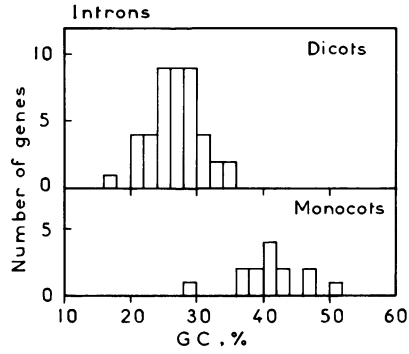

Fig. 7. The numbers of genes from dicots and monocots are plotted against the GC levels of their introns. Average values were used for introns and exons belonging to the same gene.

in the case of monocots. In spite of the fact that histograms covered now such broad ranges, a very broad maximum centered around 45% GC was still apparent in the case of dicots, whereas in the case of monocots the high GC coding sequences were further shifted to the right.

The GC distribution of coding sequences from individual genomes matched the overall distributions just discussed (not

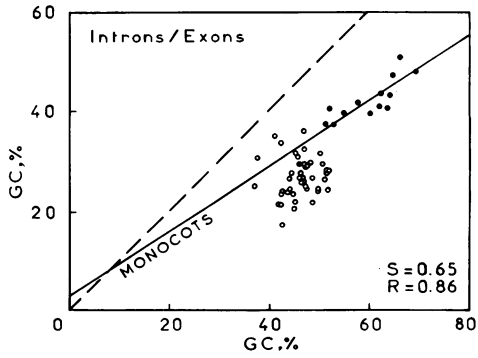

Fig. 8. GC levels of introns (ordinate) are plotted against GC levels of exons (abscissa) corresponding to the same genes from dicots (open circles) and monocots (closed circles). Average GC levels for introns and exons belonging to the same genes were used. In the case of monocots, the slope(S) and the correlation coefficient (R) are indicated. For the genes used in this plot, see Materials and Methods.

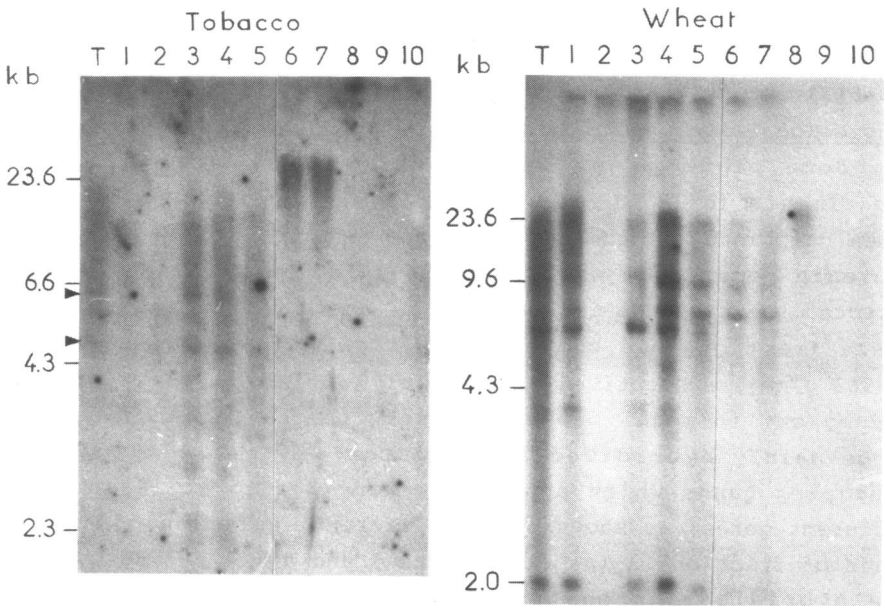

Fig. 9. Location of Cab and HMW glutenin genes in Cs_2SO_4/BAMD fractions from tobacco and wheat, respectively. 10 µg of total DNA and DNA from Cs_2SO_4/BAMD fractions in amounts corresponding to 100 µg of total DNA were processed as indicated in Materials and Methods. Arrows indicate hybridization bands.

shown). Preliminary results also indicate higher GC levels (in all codon positions) for homologous coding sequences of monocots compared to dicots, in agreement with the results of Niesbach-Klösgen et al. (20) for the coding sequences of the chalcone synthase gene.

Compositional distributions of introns

The compositional distribution of introns displayed in Fig. 7, shows higher GC values in the case of monocots compared to dicots. Plots of GC levels of introns against GC levels of exons from the same genes (Fig. 8) showed that in both dicots and monocots, GC levels of introns were remarkably lower than GC levels of exons. In the case of genes from dicots, the average GC levels of introns and exons were 27.0% (σ= 3.9%) and 46% (σ= 3.8%), respectively. In the case of exons from monocots, which covered a broad GC range (42-76%), a linear relationship was found between exons and introns from the same genes (Fig. 8) and

the difference in GC level between exons and introns also was around 20%; moreover, GC levels of introns from monocots (29-51%) were higher than those from dicots (17-36%).

Genome localization of plant genes

Some plant genes were localized by hybridization experiments in Cs_2SO_4/BAMD fractions. As shown in Fig. 9, these genes comprised (i) the cab gene from tobacco; and (ii) the HMW glutenin genes from wheat. In the first case, the wheat probe detected the cab gene in fractions 3, 4 and 5 from tobacco DNA; these fractions had buoyant densities of 1.6957, 1.6963 and 1.6963 g/cm^3, respectively; they only differed, therefore, by 0.6 mg/cm^3 in modal buoyant density. In the second case, the probe mainly hybridized on fraction 4 (ρ= 1.7029 g/cm^3), producing bands which apparently were due to at least three different genes, as shown by the different, weaker hybridization bands of fractions 3 (ρ= 1.7019 g/cm^3) and 5 (ρ= 1.7033 g/cm^3). The hybridization on fraction 1 (pellet) was due to some contamination of the pellet by GC-richer fractions (as indicated by the analytical profile of Fig. 2 and by the absence of hybridization in fraction 2; see also Salinas et al., (10), for a similar artefact in mouse DNA).

DISCUSSION

Since the present investigations were inspired by previous work on the vertebrate genome, we will first outline the conclusions reached on the latter (see refs. 4-8). Very briefly (see ref. 4, for a review), it has been shown that the vertebrate genome is made up of long (over 300 Kb), compositionally fairly homogeneous DNA segments, which were called isochores. These segments are generally GC-poor in the vast majority of cold-blooded vertebrates. In contrast, about one third of the genome is made up by GC-rich isochores in warm-blooded vertebrates. GC levels of genes are linearly related in all their codon positions to the GC levels of the isochores in which they are located; this entails not only a different codon usage and a different discrimination against CpG doublets, but also a different aminoacid composition in the encoded proteins (5). Genes are predominantly GC-poor in

cold-blooded vertebrates and show a compositional distribution paralleling that of DNA molecules (50-100 Kb in size); in contrast, genes are predominantly GC-rich in warm-blooded vertebrates, and their concentration increases in isochores of increasing GC (6,7). Since the relative amounts of isochores decrease with increasing GC, the compositional distribution of genes in warm-blooded vertebrates is almost the mirror image of that of DNA molecules. In fact, most genes from warm-blooded vertebrates underwent strong increases in GC by point mutations at the time of the transition from cold-blooded to warm-blooded vertebrates (8).

The compositional compartmentalization of plant genomes.

The first major conclusion of the present work is that the genomes of the flowering plants studied exhibit a compositional compartmentalization. Early preparative centrifugations in CsCl of DNAs from wheat (21) and from three Cucurbitaceae (1) had already shown a remarkable compositional heterogeneity in the genomes of these plants. This has been confirmed and better defined here for a larger number of plant families using a fractionation technique having a higher resolving power. What has been shown, in addition, is that the genomes investigated are made up of DNA molecules, that not only belong to classes exhibiting different GC levels, but derive from larger DNA segments, the isochores, which are compositionally homogeneous over distances at least twice the average size of DNA molecules, namely over at least 100-200 Kb.

The intramolecular compositional homogeneity over these distances is indicated by the hybridization results (Fig. 9). In fact, the very narrow buoyant density range of Cs_2SO_4/BAMD fractions hybridizing a given probe indicates that the DNA molecules, which carry the sequences detected by the probe have base compositions within less than 1% GC, irrespective of the extension of flanking segments on either side of the sequence tested. (It should be recalled that DNA molecules are produced by random fragmentation of chromosomal DNA during preparation and may carry the sequence at any position). This situation is similar to that already found in vertebrate genomes (see, for example, refs. 10,22).

Three other findings are in line with the existence of isochores in the plant genomes investigated. (i) The close GC values of genes belonging to the same gene clusters (like the genes for storage proteins of monocots) are in agreement with the compositional compartmentalization under discussion. Indeed, gene clusters are much smaller in size than isochores and are therefore expected to lie within single isochores. (ii) The limited results on gene localization already available suggest a linear relationship between the GC levels of coding sequences and those of the DNA molecules carrying them. A comparison of the compositional distributions of DNA molecules and coding sequences indicates that the latter are higher in GC than the former in both dicots and monocots. This means that intergenic non-coding sequences (which represent the vast majority of plant DNAs and practically correspond to DNA molecules) are lower in GC than coding sequences. (iii) The linear correlation between GC levels of introns and exons from the same monocot genes also fits with an isochore organization of the plant genome. The GC level of introns is, however, lower than that of corresponding exons. Incidentally, the lower GC levels of introns (and the higher GC level of exons) relative to intergenic sequences does not affect the intramolecular homogeneity over the long DNA regions described above, because of the small sizes of both introns and exons.

Different compositional patterns in the genomes of the dicots and monocots studied

The second major conclusion is that the compositional patterns of the monocots studied is remarkably different from that of dicots. Indeed, while at the DNA level the monocots studied show a higher GC (and in the case of maize and wheat, a skewness towards even higher GC values) relative to dicots, at the coding sequence level, monocots (essentially barley, maize and wheat) show a number of even more remarkable differences. (i) The majority of genes from monocots is very much shifted towards higher GC values compared to genes from dicots, (Figs. 4, 7,8); interestingly, these genes are mostly housekeeping genes, whereas tissue-specific genes, like those of seed storage proteins are GC-poor, and have GC levels close to those of most

genes from dicots. (ii) Since GC-rich genes from monocots appear to be present in GC-rich isochores which are scarce in monocots, the gene concentration in such isochores is likely to be higher than that of GC-poor genes in the more abundant GC-poor isochores; this would lead, in monocots, to a gene distribution which is the mirror image of the DNA distribution. (iii) The different compositional distributions of coding sequences from monocots relative to dicots are due to large differences in the GC distributions of third codon positions and to more moderate, yet significant, differences in those of first + second codon positions; these differences entail not only large changes in codon usage, but also aminoacid changes in the corresponding proteins.

It should be stressed that the compositional patterns of the dicots and monocots just described are astonishingly similar to those previously found for cold-blooded and warm-blooded vertebrates, respectively (4-8). Such similarity even extends to the distribution of housekeeping and tissue-specific protein genes in monocots and warm-blooded vertebrates.

The phylogenetic range of the different compositional patterns of dicots and monocots.

An obvious question at this point is how phylogenetically widespread are the differences observed between the dicots and monocots studied. This question should be analyzed at both (i) the DNA and (ii) the coding sequence level.

(i) In the case of dicots, the three species investigated belong (23) to three different orders Fabales (pea), Asterales (sunflower), Scrophulariales (tobacco), and are characterized by DNA molecules centered on GC values close to those of dicots from other orders studied so far (1). The three species investigated appear, therefore, to be good representatives for at least several orders of dicots.

In contrast, the monocots studied here belong (24) to the same order, Poales and to the same family, Poaceae or Gramineae, with maize belonging to the Zea group of the sub-family Panicoideae, rice to the Oryza group of the sub-family Bambooideae, and wheat to the Triticum group of the sub-family Pooideae. The few available GC values of DNA from monocots are

uniformly spred over a wide range (37 to 57%; 18,25). Moreover, our results indicate differences even between rice on one hand, and maize and wheat on the other. The compositional distributions found in Zea and Triticum can, therefore, be extrapolated only to the species closest to these groups.

(ii) As far as coding sequences from the bank and from the literature are concerned, data concern a number of orders in the case of dicots, whereas in the case of monocots, they practically only concern three species, (barley, maize and wheat) from two sub-families of Gramineae.

In conclusion, the comparisons made here can be said to be valid on one hand for at least several orders of dicots and on the other for at least two sub-families of the Gramineae family from monocots. The evolutionary implications of this work will be discussed elsewhere.

ACKNOWLEDGEMENTS

We thank EMBO, Heidelberg (Federal Republic of Germany), for a short-term fellowship to J.S. and the Fondazione V.V. Landi, Accademia dei Lincei, Roma, Italy, for a long-term fellowship to G.M. We also thank Drs. N.H. Chua and J.M. Malpica for the gift of DNA probes, and our colleagues Giacomo Bernardi and Jan Filipski for their help in some experiments.

*To whom correspondence should be addressed

REFERENCES

1. Ingle, J., Pearson, G.G. and Sinclair, J. (1973) Nature New. Biol. 242, 193-197.
2. Shapiro, H.S. (1976) In : Fasman, G.D. (ed.) Handbook of Biochemistry and Molecular Biology, 3rd edn, vol II, pp. 241-281, CRC Press Inc.
3. Sorenson, J.C. (1984) Advances in Genetics 22, 109-144.
4. Bernardi, G., Olofsson, B., Filipski, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M. and Rodier, F. (1985) Science 228, 953-958.
5. Bernardi, G. and Bernardi, G. (1986) J. Mol. Evol. 24, 1-11.
6. Mouchiroud, D., Fichant, G. and Bernardi, G. (1987) J. Mol. Evol. 26, 198-204.
7. Perrin, P. and Bernardi, G. (1987) J. Mol. Evol. (in press).
8. Mouchiroud, D., Gautier, C. and Bernardi, G. (1988) J. Mol. Evol. (in press).
9. Cortadas, J., Macaya, G. and Bernardi, G. (1977) Eur. J. Biochem. 76, 13-19.

10. Salinas, J., Zerial, M., Filipski, J. and Bernardi, G. (1986) Eur. J. Biochem. 160, 469-478.
11. Kislev, N. and Rubenstein, I. (1980) Plant Physiol. 66, 1140-1143.
12. Forde, J., Malpica, J.-M., Halford, N.G., Shewry, P.R., Anderson, O.D., Greene, F.C. and Milflin, B.J. (1985) Nucl. Acids Res. 13, 6817-6832.
13. Lamppa, G.K., Morelli, G. and Chua, N.-H. (1975) Mol. Cell. Biol.5, 1370-1378.
14. Gouy, M., Milleret, F., Mugnier, C., Jacobzone, M. and Gautier, C. (1984) Nucl. Acids Res. 12, 121-127.
15. Kirk, J.T.O. (1967) J. Mol. Biol. 28, 171-172.
16. Szybalski, W. and Szybalski, E. (1971) In : Cantoni, G.L. and Davidson, D.R. (Eds.) Procedures in Nucleic Acid Research, vol 2, pp. 311-354.
17. Kemp, J.D. and Sutton, D.W. (1976) Biochim. Biophys. Acta 425, 148-156.
18. Wagner, I. and Capesius, I. (1980) Biochim. Biophys. Acta 654, 52-56.
19. Altenbach, S.B., Pearson, K.W., Leung, F.W. and Sun, S.S.M. (1987) Plant Mol. Biol. 8, 239-250.
20. Niesbach-Klösgen, U., Barzen, E., Bernhardt, J., Rohde, W., Schwarz-Sommer, Zs., Reif, H.J., Wienand, U. and Saedler, H. (1987) J. Mol. Evol. 26, 213-225.
21. Huguet, T. and Jouanin, L. (1971) Biochim. Biophys. Acta 262, 431-440.
22. Zerial, M., Salinas, J., Filipski, J. and Bernardi, G. (1986) Eur. J. Biochem 160, 479-485.
23. Altman, P.L. and Dittner, D.S., (eds.) (1973) Biology data book, 2nd edn, Bethesda, Md, USA.
24. Dahlgren, R.M.T., Clifford, H.T. and Yeo, P.F. (1985) The Families of Monocotyledones, Springer-Verlag, Berlin.
25. Dokhiem, M., Sulimova, G.E. and Vanyushin, B.F. (1978) Biokhimiya 43, 1312-1317.